DOCUMENT RESUME

ED 433 351 TM 030 020

AUTHOR Aaron, Bruce C.; Parshall, Cynthia G.; Kushner, Susan N.

TITLE Supporting and Assessing Higher-Order Learning on the World Wide Web: Implementation and Evaluation of a Classroom Discussion Support System.

PUB DATE 1999-04-23

NOTE 30p.; Paper presented at the Annual Meeting of the American Educational Research Association (Montreal, Quebec, Canada, April 19-23, 1999).

PUB TYPE Reports - Research (143) -- Speeches/Meeting Papers (150)

EDRS PRICE EDRS Price MF01 Plus Postage. PC Not Available from EDRS.

DESCRIPTORS *Discussion; *Educational Assessment; Educational Research; *Graduate Students; Graduate Study; Higher Education; Learning; Pilot Projects; *Student Attitudes; *World Wide Web

IDENTIFIERS *Higher Order Learning

ABSTRACT

A pilot project demonstrated the remote use of an application designed with hypertext markup language (HTML) as an assessment and discussion support system for 28 students enrolled in a graduate educational research methods course at a midwestern state university. The web page was developed to support classroom learning and assessment by guiding students through a process designed to accomplish the following: (1) accommodate the construction of knowledge for targeted learning outcomes at higher-order cognitive levels by guiding each student through an evaluation of the responses of others; (2) assess individual and group performance on these learning objectives; and (3) measure change in student knowledge and opinion as a result of the process. Formative information from the students indicated that the interface was straightforward, that navigation was clear and uncomplicated, and that they felt engaged by the process. The effects of uncontrolled, multifaceted item stimuli derived from users' open-ended responses are considered, and the effects of the process on change in opinion and learning are examined. Three appendixes contain two sample on-line items and a sample of student opinions about the process. (Contains 2 figures, 10 tables, and 10 references.) (Author/SLD)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED 433 351

Supporting and assessing higher-order learning on the World Wide Web: Implementation and evaluation of a classroom discussion support system

Bruce C. Aaron
Andersen Consulting
Cynthia G. Parshall
Institute for Instructional Research and Practice, USF
Susan N. Kushner
University of Akron

PERMISSION TO REPRODUCE AND DISSEMINATE THIS MATERIAL IN MICROFICHE, AND IN ELECTRONIC MEDIA FOR ERIC COLLECTION SUBSCRIBERS ONLY, HAS BEEN GRANTED BY

Bruce Aaron

2A

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

☑ This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

TM030020

A paper presented at the annual meeting of the
American Educational Research Association, Montreal, April 23, 1999

BEST COPY AVAILABLE

Abstract

This pilot project demonstrates the remote use of an application designed with hypertext markup language (HTML) as an assessment and discussion support system for students enrolled in a graduate educational research methods course at a Midwestern state university. The web page was developed to support classroom learning and assessment by guiding students through a process designed to accomplish the following: (1) accommodate the construction of knowledge for targeted learning outcomes at higher-order cognitive levels by guiding each student through an evaluation of the responses of others, (2) assess individual and group performance on these learning objectives, and (3) measure change in student knowledge and/or opinion as a result of the process. Formative information provided by students indicated that the interface was straightforward, that navigation was clear and uncomplicated, and that they felt engaged by the process. Advantages and disadvantages of the application are examined. The effects of uncontrolled, multifaceted item stimuli derived from users' open-ended responses are considered. In addition, the effects of the process on change in opinion and learning are examined. Presented also are recommendations for future research and development of consensus-building discussion support systems.

Supporting and assessing higher-order learning on the World Wide Web:
Implementation and evaluation of a classroom discussion support system

As asynchronous computer conferencing is increasingly used in instructional settings, educators and instructional designers using computer conferencing in teaching are drawing on principles of collaborative learning and the social construction of knowledge (Murphy & Collins, 1997). Rapid progress in the development of hardware, software, and internet-based applications accommodates the local development of electronic environments that involve students in interactive knowledge construction. Among the outcomes of participation in such environments are problem-solving, decision-making, and product development.

Software supporting the work of groups whose members might be separated by time and/or distance is referred to as groupware (Greenberg, 1991), and the development of commercial applications such as LotusNotes© indicates the interest and demand for such software tools. However, as the current project demonstrates, local developers can construct an interactive, web-based environment to support and mediate student problem-analysis or decision-making.

Rich information for assessing the positions of participants exists naturally within this environment. The project described herein provides a framework for the measurement of learning outcomes and/or participant opinion, dependent upon whether the process is focused on collaborative decision-making or, as in the current demonstration, is focused on classroom instructional objectives.

In addition to instructional applications, such systems might be used to support decision-making in schools, school districts, or other organizations. In Florida, for example, recent educational reform and reorganization mandate systemic change in school decision-making by prescribing a decentralized, collaborative process which purports wide inclusion of community stakeholders (Florida Commission on Education Reform and Accountability, 1992). In general, such attempts to decentralize decision-making acknowledge the increasingly dynamic nature of the social and technological environments within which modern organizations operate, and recognize that optimal decision-making is a mandatory response to changing environmental circumstances (Pasmore, 1988). As a restructuring strategy, modifying decision processes resonates with Griffiths' (1959) position that the very structure of an organization is determined by the nature of the decision-making procedures it employs.

Benefits attributed to group decision-making include: (a) better, more informed decisions, (b) enhancement of participants (e.g., improved communication and group process skills), (c) greater ownership of, and responsibility for, the derived decisions, and (d) an increased sense of commitment to the goals of the organization (Owens, 1991). Group decision-making offers particular advantages in the amount of information that can be collected, the creative generation of alternatives, and the evaluation of alternatives. These activities are also important steps in problem-solving (Andre, 1986), indicative of the heuristic overlap of problem-solving and decision-making. Disadvantages in group decision-making largely ensue from the need for negotiation and compromise among group members.

In Florida, School Advisory Councils are site-based decision-making groups mandated by educational reform legislation. However, it typically has been difficult for these councils to fully represent their diverse groups of stakeholders in a participatory, decentralized decision-making process in a manner that meets the intent of the legislation. Such decision-making groups are often beleaguered by diverse obstacles, ranging from logistical challenges to the management of complex group dynamic processes (Aaron, 1992).

The importance of establishing a more participatory decision-making process in public education might extend far beyond the need to satisfy the regulatory mandates of reform. Gerzon (1997) argued that the very survival of public education depends on involving its broadening constituency in active consideration of the conflicting perspectives of the diverse groups it comprises. Computer-assisted group decision support systems might be one method of establishing a decentralized, collaborative, inclusive, and accountable process for making decisions about schools and schooling.

Purpose

The purpose of this project was to develop, pilot, demonstrate, and refine a process and an accompanying HTML-based computer interface that would accomplish the following: (1) assess individual and group performance on learning outcomes of interest at higher-order cognitive levels, (2) accommodate the construction of new knowledge for those outcomes by guiding each student through structured evaluation of the responses of other students, and (3) measure change in student knowledge

and/or opinion as a result of the process. This process is conceived as generalizable to a range of group problem-solving and decision-making situations.

Method

The project used a remote webpage located on a server at a large Southeastern state university as an assessment and discussion support system for students enrolled in a graduate educational research methods course at a Midwestern state university. In using the system, a student does the following: (1) responds with Likert-type responses to presented problem scenarios that are based on the instructional objectives of the class, (2) provides open-ended elaborations explaining his/her chosen position on the Likert-type scale, (3) reads, and offers Likert-type responses to, the open-ended elaborations of other students, and finally, (4) re-analyzes each scenario and restates his/her position on the Likert-type scale (see appendices 1 and 2 for examples of each step). Students must therefore analyze and evaluate the components of the presenting problem scenarios and other students' responses to the problems, and synthesize this information in formulating their final answer or opinion. Thus the system promotes and assesses learning at higher levels that correspond, for example, to Bloom, Madaus, and Hastings' (1981) levels of analysis, synthesis, and evaluation.

In addition, students respond to teacher-embedded elaborations, which appear on-screen as the ostensible open-ended elaborations of other participants, providing a means of assessing students' knowledge of particular concepts and ideas representing specific targets of instruction. This allows a mapping of students' conceptual foundations regarding particular instructional objectives of interest.

The process supported by the system encompassed two consecutive stages. Students were given two corresponding windows of opportunity within which to respond. Responses were asynchronous within each stage. During stage 1, students logged on, received directions, and responded to items. Items consisted of open-ended scenarios depicting situations relevant to educational research, and revealing a particular course of action or point of view (see appendix 1 for a sample item). Students responded first by indicating the extent of their agreement or disagreement with a particular course of action or point of view, and then explained their professed position with a short elaboration, entered into a scrollable text field. These elaborations were stored in a database keyed to their originator (at this point in the process, these

responses might be scored globally or analytically according to an instructor's rubric, to provide a higher-order summative measure of performance on the instructional objectives reflected in the scenarios).

During stage 2, students were presented with the elaborations of others, and asked to indicate the extent of their agreement with each expressed opinion or answer to the presented scenario (see appendix 2). Since a participant can learn from or be persuaded by other students' responses, stage 2 concluded with a re-assessment of the student's position on the action or point of view presented in each scenario, in order to measure any change in response. Finally, students in these trials responded on-line to formative questions regarding use of the system. The primary project used three iterations, or trials, of this two-stage process over the course of one semester. A smaller replication is also reported for one trial of the two-stage process, conducted during the following semester.

The page was designed using the Netscape 3.0 Gold© editor. Microsoft Excel© was use to manage the data. CGI scripts were used to interface the page/server and database.

Data Analysis

All forced-choice responses were translated to a 4-point, Likert-type scale representing level of support for, or agreement with, the item stimuli. These scale responses represented support for particular points of view for certain of the presented problem scenarios, or the traditional "Strongly Disagree (1)", "Disagree (2)", "Agree (3)", and "Strongly Agree (4)" positions for registering opinion about the elaborations of other students.

Expected scale values for responses to the elaborations of others were derived using the simple logical model displayed in table 1. If, for example, subject A, who is in strong disagreement with a certain idea, offers an elaboration of that disagreement to subject B, who is conversely in strong agreement with the idea, then subject B could be predicted to disagree strongly with subject A's elaboration. Residuals were calculated as the difference between an individual's observed and expected response to an elaboration. Positive residuals indicate greater-than-expected support, while negative residuals indicate unexpected disagreement. Expected and residual values were used to identify consistency and relationships among student responses.

Table 1: Model for expected values of responses to elaborations

			Scale Value of Person Responding to Elaboration			
			Strongly Disagree	*Disagree*	*Agree*	*Strongly Agree*
			1	*2*	*3*	*4*
Scale Value for Elaboration	*SD*	*1*	**4**	**3**	**2**	**1**
	D	*2*	**3**	**4**	**2**	**1**
	A	*3*	**1**	**2**	**4**	**3**
	SA	*4*	**1**	**2**	**3**	**4**

Simple summary statistics, based on rating scale values, were calculated to describe the group's position and variability for each problem presented, and to indicate change in position between stages 1 and 2.

Participants

Primary participants in the study were 28 graduate students enrolled in a required research methods course in the College of Education at a large metropolitan university in the Midwest. Results are also reported for a replication conducted with another group of 24 student similarly enrolled during the following semester. All but five of the primary participants were part-time students and most held full-time jobs (primarily teaching positions). This particular section of the course was being offered in a distance education format. Ten students attended class at the university site and the remaining 18 students attended class at the distance education site, approximately 25 miles from the campus. Each of the classrooms was identically well-equipped with modern distance education technology, including an instructor camera, two student cameras, a document camera, a video projector and monitors, and electronic whiteboard technology (the latter allows both student and teacher, who are separated geographically, to write on the same board).

Half of the students reported that they were just beginning their graduate studies. Although a few students previously had taken a research-related course, 80 percent had no prior course experience in research methods. Approximately three-fourths of the students reported that they had at least some knowledge of internet applications (e.g., email and the world-wide web). Slightly less (two-thirds) reported

that they had at least minimal skills and experience using these applications. Although three-fourths of the students had home computers, slightly less than half had internet access from home. The remaining students accessed the internet from university computer labs, other public access labs (e.g., the public library), or from work. Many were teachers who had internet access in their classrooms.

Results

Student Self-Reported Characteristics

Upon completing each exercise, students were asked general questions that included inquiry about their comfort with use of the computer and internet, their location, and the type of web browser they were using. Responses are displayed in tables 2 through 5. Most notable in these data are apparent trends toward more positive self-efficacy in using computers, and the world wide web in particular, over the course of the exercise (see tables 4 and 5).

Table 2. Locations from which the exercises were completed.

	TRIAL 1		*TRIAL 2*		*TRIAL 3*	
Location	*f*	%	*f*	%	*f*	%
Home	9	36	7	28	9	35
Work	5	20	6	24	10	38
University	1	4	2	8	1	4
Distance Site	9	36	8	32	5	19
Home of Friend/Relative	1	4	2	8	1	4
TOTAL	25	100	25	100	26	100

Table 3. Internet browsers used

	TRIAL 1		TRIAL 2		TRIAL 3	
Browser or Service	*f*	%	*f*	%	*f*	%
America Online©	2	8	2	8	2	8
Internet Explorer©	2	8	4	16	4	15
Netscape©	20	80	19	76	19	73
Prodigy©	1	4			1	4
TOTAL	25	100	25	100	26	100

Table 4. Self-rating of computer skill.

"How would you rate your computer skills?"						
	TRIAL 1		TRIAL 2		TRIAL 3	
	f	%	*f*	%	*f*	%
Beginner	11	44	10	40	7	27
Intermediate	13	52	14	56	17	65
Advanced	1	4	1	4	2	8
TOTAL	25	100	25	100	26	100

Table 5. Self-rated World Wide Web skill.

"How would you rate your skills using the World Wide Web?						
	TRIAL 1		TRIAL 2		TRIAL 3	
	f	%	*f*	%	*f*	%
Beginner	18	72	16	64	12	46
Intermediate	7	28	8	32	13	50
Advanced			1	4	1	4
TOTAL	25	100	25	100	26	100

Design and Operation

Formative information provided by students indicated that the interface was straightforward, that navigation was clear and uncomplicated, and that they felt engaged by the process. Based on student comments, important advantages of the process seem to be the extent to which students become engaged in evaluating the content matter at a higher-order level. Users reported being particularly interested in evaluating the responses of other students, and noted that the process encouraged them to carefully consider their opinions and the positions and ideas of others. The anonymous nature of the interface apparently contributes to a focused exchange and evaluation of ideas.

Users were also asked how comfortable they would be taking a test in this web-based format. Responses to this question are summarized in table 6. At each trial, about 80% of students said that they would be "very comfortable" or "somewhat comfortable" being assessed within this system. The proportion of students indicating that they would be "very comfortable" increased from 20% to 50% over the course of the semester.

Table 6. Level of comfort with Web-based assessment

"How comfortable do you think you would be taking a test in this type of format?"	*TRIAL 1*		*TRIAL 2*		*TRIAL 3*	
	f	%	*f*	%	*f*	%
Very Uncomfortable	2	8	3	12	1	4
A Little Uncomfortable	3	12	2	8	4	15
Somewhat Comfortable	15	60	13	52	8	31
Very Comfortable	5	20	7	28	13	50
TOTAL	25	100	25	100	25	100

Assessing Student Responses

As a summative measure, students' elaborations, which expressed and defended the positions taken relative to the problem presented, were available to the instructor to be scored as essay responses to provide performance assessment of the instructional objectives reflected in the scenarios. In addition, a primary purpose of the system is to

accommodate the construction of new knowledge at higher-order cognitive levels by guiding each student through an evaluation of the ideas of others. Of interest, therefore, was the extent to which students' conceptions were modified by the exercise. Table 7 displays for each trial mean rating scale data for each problem scenario for Stage 1 (before reviewing other students' elaborations) and Stage 2 (after considering and responding to other students' elaborations), mean change scores, and effect size estimates for change.

Table 7: Summary of student response means

	Rating Scale Scores						Effect size (*d*)
	Stage 1		Stage 2		Change		
	Mean	SD	Mean	SD	Mean	SD	
Trial 1							
Scenario 1	2.86	.79	3.05	.59	.19	.98	.24
Scenario 2	3.67	.48	3.78	.44	.09	.44	.23
Scenario 3	3.19	.75	3.52	.68	.33	.58	.44
Trial 2							
Scenario 1	2.56	1.04	2.62	.90	.04	.81	.06
Scenario 2	3.64	.57	3.27	.67	-.40	.76	-.65
Trial 3							
Scenario 1	1.79	.66	1.83	.70	.04	.91	.06
Scenario 2	2.00	.66	1.96	.46	-.04	.55	-.06

A replication of the process was conducted in the course during the following semester with a new group of students (n=24), using the problem scenario that had elicited greatest change in student conceptions (trial 2, scenario 2, effect size = -.65). Figure 1 displays this problem.

Figure 1. Trial 2, Scenario 2

2.	Tony Fernwood and Marcus Grizson are planning to use survey methods to collect data on the recycling habits of residents of Summit County. Tony and Marcus agree to draw a random sample of residents from the county to participate in their study. They are torn between using a telephone interview or a mailed questionnaire to collect the data. Which method would you advise Tony and Marcus to use? O Definitely use a telephone interview O Probably use a telephone interview O Probably use a mailed questionnaire O Definitely use a mailed questionnaire Please explain your position in the space below:
	[elaboration is entered into scrollable text field]

Using this scenario, the process worked similarly when applied to the new sample of students. In the semester 2 replication, group position during the phone-versus-mail survey problem demonstrated similar substantial change. Table 8 displays mean rating scale values for this problem scenario at Stage 1 and Stage 2, mean change scores, and effect size estimates for change, for each semester.

Table 8: Summary of group response across semesters for phone versus mail survey problem

Problem Scenario: Phone Survey versus Mail Questionnaire	*Rating Scale Scores*						Effect size (*d*)
	Stage 1		Stage 2		Change		
	Mean	SD	Mean	SD	Mean	SD	
Semester 1	3.64	.57	3.27	.67	-.40	.76	-.65
Semester 2	3.69	.47	3.42	.58	-.40	1.04	-.76

Student Feedback

An open-ended question asked students to describe strengths or weaknesses of the activity, the extent to which they felt challenged, and/or suggestions for improvement. Across the two semesters, 153 comments were collected and categorized. Eight categories of comment accounted for 75% of the total. Comments were designated as positive, negative, or neutral in implication, with respective values of 1, –1, and zero assigned to each. The sum of these values across comments was 51, with a mean of .33.

The predominant observation provided by students (comprising about 37% of comments) referred to the interest or value in evaluating, rating, and reflecting on the responses of others. Some noted that the anonymous and self-paced nature of the activity contributed to the value of this reflection. This category of comment is dubbed Reflective Evaluation in the table of comments in Appendix 3.

About 8% of comments indicated that the activity was not particularly challenging. On the other hand, about 4% noted that the activity was challenging. Other noteworthy positive comments referred to the convenience of doing class exercises via the Internet, and general observations that the activity was interesting. Other noteworthy negative comments observed that the design of the process was unnecessarily repetitious (i.e., included evaluation of redundant comments), that feedback regarding the correct solution to the presented problem should have been provided, or cited various technological or mechanical difficulties that were external to the instructional design of the exercise. Categories of comments, and their relative frequencies, are listed in Appendix 3 along with exemplars.

Discussion

Assessing Student Elaborations

Since the system is designed to support group processes of discussion, decision-making, and problem-solving, it must use the open-ended responses of group members as item stimuli, a prospect presenting unique advantages and disadvantages. Within this system, the extent to which users' open-ended responses (here termed elaborations) are pertinent to the question or problem being represented in the prompt can vary unpredictably. In this open-ended environment, as in the process of human discussion and deliberation, these elaborations are stimuli for other responses, but these item

stimuli are necessarily multifaceted and uncontrolled. This system attempts to measure user response to these elaborations as a means of assessing learning and/or opinion. This seems problematic in that psychometric accuracy traditionally is predicated upon careful development and control of the items to which users respond. Nonetheless, the data collected in using the system can provide useful estimates derived from anomalous responses.

This is apparent when users disagree with elaborations with which they would be predicted to agree, and vice versa. These anomalies, or residuals, can be used purposively to guide and support the decision-making or learning process. These residuals can be conceived as resulting from two general processes: a) error, or b) persuasion (which can be conceived as positive or negative).

In these data, error was apparent in inconsistencies that probably resulted from users' coding errors, particularly evident for users at the extreme ends of the scale of opinion who likely reversed the response values. Evidence of persuasion is found in the accumulation of large residuals when users tend to agree or disagree with a particular elaboration to a larger extent than would be predicted by their own stated positions. Thus, these residual indices can be used to identify elaborations that are useful in building consensus (when the system is used to support group decision-making), or to identify where new learning is occurring (when the system is being used to support and assess learning).

As an example, consider the phone-versus-mail survey problem (displayed in figure 1), by which the process demonstrated the largest change in group position. This effect was replicated across two semesters with different classes. Scale values 4 through 1, respectively, were assigned to the responses selected for this scenario, indicating strength of preference for using telephone interviewing rather than mail surveys. Two particular elaborations, and the scale values associated with them, are presented in table 9. In each semester, student responses to each of these elaborations resulted in particularly high positive discrepancies between the observed and predicted levels of agreement.

Table 9. Two elaborations For Scenario 2, Trial 2

Elaboration #	*Scale Value*	*Elaboration*
6	1	*Mailed questionnaires can be a very effective way to collect data. By writing an appropriate cover letter, including an incentive, and making contacts, Marcus and Tony can get a satisfactory return rate.*
7	2	*Marcus and Tony could include a copy of their questionnaire in the monthly service bills that residents of Summit County receive. Then the questionnaire would look less like junk mail. Participants could get a small rebate on their bill for completing the questionnaire. This would increase the response rate.*

Presented in table 10 are item statistics for responses to these elaborations across both semesters. As indicated by the mean rating scale scores (3.64 and 3.69) on the original item (column *a*), student support for the use of telephone interviews rather than mailed questionnaires was strong, tending to predict disagreement with comments such as those displayed in table 9, which favor the use of mailed questionnaires over telephone interviews. Accordingly, as shown in table 10, the mean expected values of agreement with these elaborations were low, at 1.36 and 1.40, respectively during semester 1, and 1.31 during semester 2 (column *b*). However, the average level of agreement with these elaborations was higher, as seen during semester 1 in high mean residual values of 1.14 and 1.55, for elaborations 6 and 7, respectively, and in values of 1.17 and 1.30 during semester 2 (column *c*).

Table 10. Item statistics for group response to elaborations 6 and 7

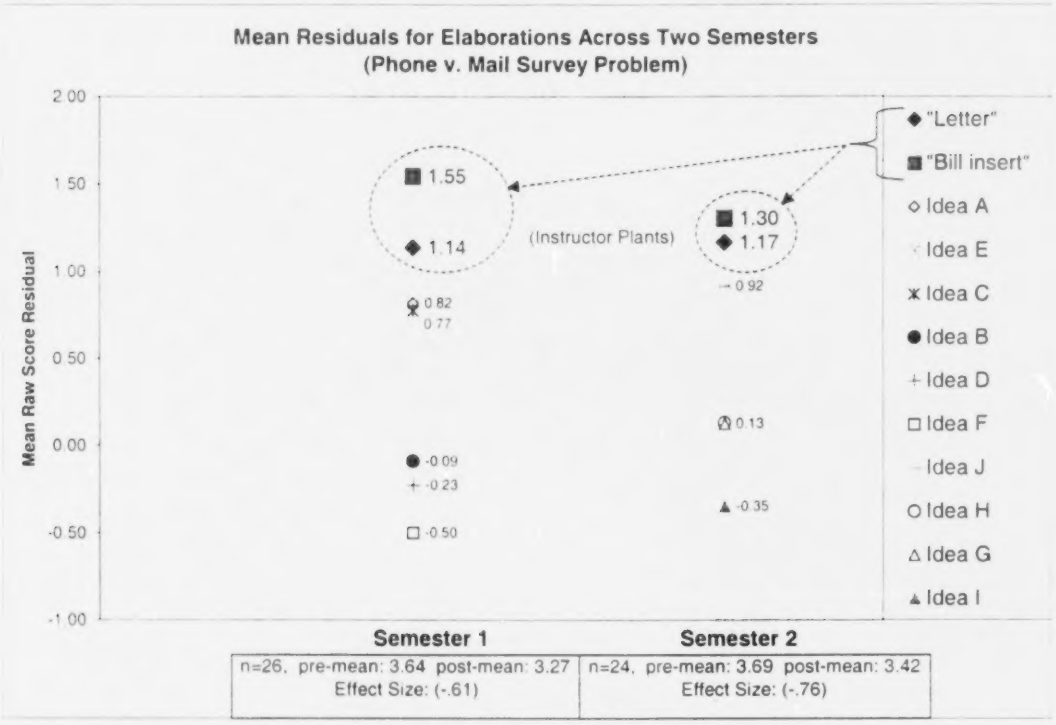
		a.	*b.*		*c.*	
		Rating Scale Value	*Expected Value of Response to Elaboration*		*Residual for Responses to Elaboration*	
Semester			Elab. 6	Elab. 7	Elab. 6	Elab. 7
1	Mean	3.64	1.36	1.40	1.14	1.55
	SD	.57	.57	.71	.56	.96
2	Mean	3.69	1.31	1.31	1.17	1.30
	SD	.47	.47	.47	.57	.82

BEST COPY AVAILABLE

Thus, students in both samples responded more positively to these comments than would be expected, providing evidence and measurement of a similar persuasive process at work in each case. Examination of the elaborations in table 9 suggests that although students were generally opposed to the positions of the elaborations' authors, they responded with unexpected support for the concrete suggestions offered in these elaborations. In this manner, item analysis information can be used to identify and present ideas that work to build consensus (when the system is used to support group decision-making), or to identify shifts in learners' schemata (when the system is being used to support and assess learning).

Notably, the elaborations displayed in table 9 were planted by the instructor. In this manner, the system can provide information not only about students' original knowledge and positions, but also can indicate where new learning (i.e. the acceptance of a particularly constructed point of instruction) has occurred. As noted previously, (see table 8) the effect sizes for mean change in position for this particular problem scenario were -.65 during semesters 1 and 2, and -.76 during semester 2 (medium-to-large effects). For each group of students, final positions were shifted toward greater agreement with the mailed questionnaire strategy, and points of view deliberately planted by the instructor were measured as having the largest contribution to persuasion. The persuasiveness of each exchanged idea, expressed as the mean scale score residual for each elaboration, is shown in Figure 2. This plot illustrates how separate samples of students convened upon the same problem found these particular ideas planted anonymously by the instructor similarly and consistently persuasive, contributing to similar substantial change in the groups' position.

Figure 2. Mean Residuals, Phone v. Mail Survey Problem



In addition to identifying particularly persuasive ideas, an examination of the data indicated other situations that were similarly identified by large positive residuals in group responses to elaborations. For example, if an elaboration specifically addressed advantages and/or disadvantages of both positions, responses to the remark tended to be more positive than predicted. These "middle-of-the road" elaborations might also be useful in developing consensus in a decision-support system.

Large negative residuals, on the other hand, identify elaborations that elicit less than expected agreement. These remarks would be expected to be relatively ineffective in developing consensus, since they are generally rejected by the group to an unexpected extent, regardless of group members' individual positions. An examination of items with large negative residuals suggested that many of these elaborations simply contained specific points which most users seemed relatively hard-pressed to support. In addition, incoherence (i.e., content relatively unrelated to the prompt) in an elaboration also contributed to large negative residuals.

Other explanations for these discrepancies are possible. For example, incongruity between a respondent's rating scale score and elaboration can contribute to the accumulation of residuals across users' responses to the elaboration. In these data, this particular inconsistency was found for some cases, suggesting the need for revision of the item and/or screen elements to minimize users' entry errors in registering their scale responses. Such revisions might involve reversing or redesigning the scale, revising instructions, and/or including a short tutorial for using the system and responding to items.

Implications

The project suggested interesting potential for measuring knowledge (in the instructionally-focused setting) or opinion (in a decision-making setting), since information exists to scale both items and persons. The process provides a structure for facilitating and documenting problem-solving and decision-making at the individual and group level by eliciting, juxtaposing, and quantifying the solutions or opinions of participants relative to some problem. It was demonstrated that changes in individual and group positions can be shifted toward greater agreement with a targeted solution (e.g. the strategy of using mailed questionnaires) in an instructionally deliberate process, and that the extent of change, and the information related to this change, can be documented for the individual and group.

Future Research and Development

The richness of the information generated within this process suggests that an empirical, probabilistic, and objective measurement model can be successfully employed. As the measurement model is refined, it will help address an additional concern -- the amount of data and information that must be managed as the number of users and responses proliferates. It quickly becomes unreasonably lengthy, for example, to present all participant elaborations to each participant for evaluation. But with the scalability of each participant (given their rating scale responses to prompts) and each elaboration (given the scale value of the elaborator, and the scale values of responses to

the elaboration), the system might be programmed to construct a sample of responses for presentation to each participant based on logical empirical criteria. This study suggested that residuals can be useful for managing this flow of information. This potential will serve well future versions of the system, which will emphasize group decision-making and consensus-building, since participants at various points along the scale of opinion can be directed to one another according to the selection criteria which prove best at maximizing opinion change in the development of consensus within the group.

References

Aaron, B.C. (1992). Restructuring with technology: A computer-assisted model for collaborative school decision-making. Unpublished manuscript.

Andre, T. (1986). Problem solving and education. In G.D. Phye and T. Andre (Eds.) Cognitive classroom learning: Understanding, thinking, and problem-solving. (pp.169-204). New York: Academic Press, Inc.

Bloom, B.S., Madaus, G.F., and Hastings, J.T. (1981). Evaluation to improve learning. New York: McGraw-Hill Book Company.

Florida Commission on Education Reform and Accountability (1992). Blueprint 2000: A System of School Improvement and Accountability.

Gerzon, M. (1997). Teaching democracy by doing it. Educational Leadership, 54(5), 6-11.

Greenberg, S. (1991). Computer-supported cooperative work and groupware. In S. Greenberg (Ed.), Computer-supported cooperative work and groupware (pp. 1-10). San Diego, CA: Academic Press, Inc.

Griffiths, D. (1959). Administrative Theory. New York: Appleton-Centurey-Crofts.

Murphy, K.L., & Collins, M.P. (1997). Development of communication conventions in instructional electronic chats. Paper presented at the Annual Meeting of the American Educational Research Association, Chicago, Illinois, March, 1997.

Owens, R. (1991). Organizational Behavior in Education. New Jersey: Prentice-Hall.

Pasmore, W. (1988). Designing Effective Organizations: The Sociotechnical Systems Perspective. New York: Wiley & Sons.

Appendix 1

Sample on-line item, stage 1

Instructions for Web Practice Exercise #2 - Stage 1

Each of the items below presents a statement or situation relevant to educational research. For each of these items, indicate the extent to which you agree or disagree by selecting one of the responses provided. Then you will be asked to explain why you tend to agree or disagree with the statement by providing a short elaboration in the space provided.

1. Sandra Alamo is a physical education teacher at Jacobs Middle School. Sandra is a little concerned about her students' behavior in the classes that she teaches. Currently, her students are not required to wear any specific type of uniform during their physical education class, although they are required to change into activity clothing (i.e., shorts, sweat pants, tennis shoes, etc.). Sandra is considering a study to determine if wearing uniforms (white tee shirts, blue shorts, and red socks) will have an effect on students' behavior. She plans to use several of her physical education classes as her sample. Pat Williams, a colleague of Sandra's, points out that she would be better off using a simple random sample of classes throughout the school district.

What do you think Sandra should do?

- O Definitely use her own classes
- O Probably use her own classes
- O Probably use a simple random sample of classes in the school district
- O Definitely use a simple random sample of classes in the school district

Please explain your position in the space below.

Appendix 2

Sample on-line item, stage 2

Instructions for Web Practice Exercise #2 - Stage 2

This stage of the on-line web practice exercise is similar to the first, in that you will be asked to read and reflect on the three research scenarios presented during the first assessment session. During this session, however, you will be presented with other students' responses, which follow each scenario. Please examine each of these responses to the situation, and then mark the extent to which you agree or disagree with each of them. For each scenario presented below, do the following:

1. Re-read the research scenario
2. Consider various responses that other students submitted
3. Indicate the extent to which you agree or disagree with each response

The web practice exercise is organized in the following sections:

Scenario 1
Scenario 2
Background Questions

Scenario 1

Please read Scenario 1 :

Sandra Alamo is a physical education teacher at Jacobs Middle School. Sandra is a little concerned about her students' behavior in the classes that she teaches. Currently, her students are not required to wear any specific type of uniform during their physical education class, although they are required to change into activity clothing (i.e., shorts, sweat pants, tennis shoes, etc.). Sandra is considering a study to determine if wearing uniforms (white tee shirts, blue shorts, and red socks) will have an effect on students' behavior. She plans to use several of her physical education classes as her sample. Pat Williams, a colleague of Sandra's, points out that she would be better off using a simple random sample of classes throughout the school district.

Below are ten (10) separate responses to the situation described above. After reading each response, please indicate the extent to which you agree or disagree with it.

Response 1 (of 10).

" I would say that Sandra should probably use her own classes. She already knows what their attitudes and behaviors are. She would then be able to compare their attitudes and behaviors after wearing a uniform. If she were to study other schools in the district she

may get different results simply because she does not know the other students and their behaviors."

To what extent do you agree with this response to Scenario 1?

- O Strongly Disagree
- O Disagree
- O Agree
- O Strongly Agree

Review Scenario 1.

Response 2 (of 10).

" She should probably use a random sample of classes in the district to reduce the possibility of bias, or the observer having an effect on the study."

To what extent do you agree with this response to Scenario 1?

- O Strongly Disagree
- O Disagree
- O Agree
- O Strongly Agree

Review Scenario 1.

. . . . . .

[student continues scrolling through responses 3 through 9]

. . . . . .

Response 10 (of 10).

" Sandra should definitely use her own classes. Pat probably thinks that using your own classes is too much like a convenient or accidental sample, but it could also be argued that Sandra is using a purposive sample. If she uses her own classes she will have to describe her students in enough detail so other physical education teachers can decide for themselves if Sandra's students are similar to theirs."

To what extent do you agree with this response to Scenario 1?

- O Strongly Disagree
- O Disagree
- O Agree
- O Strongly Agree

Review Scenario 1.

Your Opinion.

At this point, you have reacted to other students' responses to Scenario 1. Now please re-consider Scenario 1 and select the option that best expresses your opinion.

Sandra Alamo is a physical education teacher at Jacobs Middle School. Sandra is a little concerned about her students' behavior in the classes that she teaches. Currently, her students are not required to wear any specific type of uniform during their physical education class, although they are required to change into activity clothing (i.e., shorts, sweat pants, tennis shoes, etc.). Sandra is considering a study to determine if wearing uniforms (white tee shirts, blue shorts, and red socks) will have an effect on students' behavior. She plans to use several of her physical education classes as her sample. Pat Williams, a colleague of Sandra's, points out that she would be better off using a simple random sample of classes throughout the school district.

What do you think Sandra should do?

- O Definitely use her own classes
- O Probably use her own classes
- O Probably use a simple random sample of classes in the school district
- O Definitely use a simple random sample of classes in the school district

Appendix 3

Student comments regarding use of the system

"What comments, observations, or suggestions would you like to share about this activity? Perhaps you would like to comment on the following:

- *What did you like best?*
- *What did you like least?*
- *Did you feel challenged?*
- *How might this be improved?*

Category	*%*	*Examples of Comments*
Reflective Evaluation	37.3	"... I also found myself questioning my original responses as I completed the tasks. It really made me ponder the reasons for taking the positions I chose to take." "I love doing this type of homework exercise. I like to read my classmates' answers because I can get different perspectives on the topic. It's challenging because I'm given thoughts and ideas from other intelligent people to consider. It also makes you take a stand on your beliefs, and I like having to do that." "I like being able to answer the questions without feeling rushed or that someone is looking over my shoulder." "I liked being able to 'agree or disagree' with other student responses. Unfortunately, I am not a very 'secure' student at this point and it helped to see that I had answered part one in a manner similar to other students in class..." "I liked seeing other students responses. It seemed as if I agreed with opposite responses, sometimes, because they made a valid point." "I like to read my classmates' answers because I can get different perspectives on the topic. It's challenging because I'm given thoughts and ideas from other intelligent people to consider. It also makes you take a stand on your beliefs, and I like having to do that." "I enjoy reading others' views of the scenarios, it helps to develop my own opinion...either by solidifying my responses or seeing a new 'light'." "I really liked this activity. I liked being opinionated and spouting off my view. I also enjoyed being able to critique the comments of others. I would like to keep this going. I want to see the responses to the responses. It's getting real fun now."

Category	*%*	*Examples of Comments*
Not Challenging	*7.8*	"I did not sense a great deal of challenge when completing the exercise. Giving my opinion is not a challenge."
Convenient	*6.5*	"I like the ability to complete assignments in my home at my leisure. I was able to be home with my daughter and while she slept I finished this."
Repetitious	*6.5*	"I feel that having to read 8 responses was a bit much. Most of them were very similar and it seemed redundant to have to read through all 8 I feel 5 different responses would have been enough."
Mechanical/ Technology Problems	*4.6*	"I did not like the fact that I could not print out my answers." "I still struggle with reading things for long periods of time on the screen. (I must be getting old.)" "Sometimes the online server was not usable and it caused my response to be delayed."
Interesting in General	*4.6*	"I thought it was a worthwhile activity and an interesting way in which to respond to questions."
Challenging	*3.9*	"I did feel challenged by this exercise. . ."
Lacking Feedback	*3.3*	"As I read the other student's thoughts, I realized that there were several valid approaches to the problem. What I would like to see, however, is what is the BEST answer to the questions posed by the scenarios. I think this would help me learn."

BEST COPY AVAILABLE